Semidefinite Programs for Randomness Extractors

Mario Berta[1], Omar Fawzi[2,3], and Volkher B. Scholz[4]

1 **Institute for Quantum Information and Matter, Caltech**
Pasadena, CA 91125, USA
2 **Department of Computing and Mathematical Sciences, Caltech**
Pasadena, CA 91125, USA
3 **LIP*, École Normale Supérieure de Lyon**
Lyon, 69007, France
4 **Institute for Theoretical Physics, ETH Zürich**
Zürich, 8093, Switzerland

Abstract

Randomness extractors are an important building block for classical and quantum cryptography. However, for many applications it is crucial that the extractors are quantum-proof, i.e., that they work even in the presence of quantum adversaries. In general, quantum-proof extractors are poorly understood and we would like to argue that in the same way as Bell inequalities (multi prover games) and communication complexity, the setting of randomness extractors provides a operationally useful framework for studying the power and limitations of a quantum memory compared to a classical one.

We start by recalling how to phrase the extractor property as a quadratic program with linear constraints. We then construct a semidefinite programming (SDP) relaxation for this program that is tight for some extractor constructions. Moreover, we show that this SDP relaxation is even sufficient to certify quantum-proof extractors. This gives a unifying approach to understand the stability properties of extractors against quantum adversaries. Finally, we analyze the limitations of this SDP relaxation.

1998 ACM Subject Classification E.4 Coding and Information Theory

Keywords and phrases Randomness Extractors, Quantum adversaries, Semidefinite programs

Digital Object Identifier 10.4230/LIPIcs.TQC.2015.73

1 Introduction

1.1 Randomness extractors

A randomness extractor is a procedure to distill from a weakly random system as much (almost) uniform random bits as possible. Such objects are essential in many cryptographic protocols, in particular in quantum key distribution and device independent randomness expansion [3, 26, 12, 24, 35]. In this context, the process of transforming a partly private string into one that is almost uniformly random from the adversary's point of view is called privacy amplification [5, 4]. Even though we take a cryptographic point of view in this paper, we should mention that randomness extractors are very useful combinatorial objects in particular in the study of the computational power of randomness (see [34] for a survey).

More precisely, a randomness extractor is described by a family of functions $\mathrm{Ext} = \{f_s\}_{s\in D}$ where $f_s : N \to M$. We use $N = 2^n$ to denote the input system (consisting of strings of n

* UMR 5668 LIP – ENS Lyon – CNRS – UCBL – INRIA, Université de Lyon

10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015).
Editors: Salman Beigi and Robert König; pp. 73–91
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

bits), $M = 2^m$ (bit-strings of length m) to denote the output system, and $D = 2^d$ (d bits) to denote the seed system that labels the functions f_s. Note that in a slight abuse of notation, we use the same letter for the actual set of inputs/outputs as well as its size. We say that Ext is a (k, ϵ)-extractor if for any random variable X taking values in N,

$$H_{\min}(X) := -\log p_{\text{guess}}(X) \geq k \quad \Longrightarrow \quad f_{U_D}(X) \text{ is } \epsilon\text{-close to } U_M \ , \tag{1}$$

where U_D is uniformly distributed on D and independent of X and U_M denotes the uniform distribution over M. As mentioned in the equation, the min-entropy $H_{\min}(X)$ is defined by the maximum probability of success in guessing a source X with only the knowledge of the distribution p of X. In this case, we simply have $H_{\min}(X) = -\log \max p(x)$. To quantify the distance between distributions, we use the total variation distance.[1] Equation (1) can thus be more explicitly written as

$$\forall x \in N,\ p(x) \leq 2^{-k} \quad \Longrightarrow \quad \frac{1}{D} \sum_{\substack{s \in D \\ y \in M}} \left| \sum_{x : f_s(x) = y} p(x) - \frac{1}{M} \right| \leq \epsilon \ . \tag{2}$$

Even though the concept was already present in [5, 4], the definition of randomness extractors was formulated in [23]. The typical example of a family $\{f_s\}_s$ of functions that satisfy this condition are randomly chosen functions. In fact, one can show [29, 25] that choosing D functions f_s independently at random among all the functions from N to M satisfies equation (2) with the following parameters

$$m = k - 2\log(1/\epsilon) - O(1) \quad \text{and} \quad d = \log(n-k) + 2\log(1/\epsilon) + O(1) \ . \tag{3}$$

In fact, we even know that these parameters cannot be improved except for additive constants [25]. Probabilistic constructions are interesting, but for applications we usually want the functions f_s to be efficiently computable. The most famous example of an explicit extractor is given by two-universal hash functions [5, 4, 17]. However, this construction has a seed size d that of the order of n, very far from the $\log n$ achieved by probabilistic constructions (3). Constructing efficiently computable extractors that match the parameters of randomly chosen functions has been the subject of a large body of research. Starting with the work of Nisan and Ta-Shma [22] and followed by Trevisan's breakthrough result [33], there has been a lot of progress in achieving polylogarithmic seed size, and there are now many intricate constructions that come close to the parameters in (3) (see the review articles [28, 34]).

1.2 Quantum-proof randomness extractors

For applications in classical and quantum cryptography (see, e.g., [26, 20]) and for constructing device independent randomness amplification and expansion schemes (see, e.g., [11, 21, 13]) it is important to find out if extractor constructions also work when the input source is correlated to another (possibly quantum) system Q. That is, we would like that for all classical-quantum input density matrices $\rho_{QN} = \sum_{x \in N} \rho(x) \otimes |x\rangle\langle x|$ acting on QN with conditional min-entropy

$$H_{\min}(N|Q)_\rho := -\log p_{\text{guess}}(N|Q)_\rho \geq k \ , \tag{4}$$

[1] It is more convenient here to use simply the ℓ_1 norm between the distributions, ignoring the $\frac{1}{2}$ factor in the usual definition of the total variation distance.

where $p_{\text{guess}}(N|Q)$ denotes the maximal probability of guessing the system N given Q, the output is uniform and independent of Q,[2]

$$\frac{1}{D}\sum_{\substack{s\in D\\ y\in M}}\left\|\sum_{x:f_s(x)=y}\rho(x)-\frac{1}{M}\sum_{x\in N}\rho(x)\right\|_1\leq\epsilon\ . \tag{5}$$

As observed in [19, Proposition 1], if we restrict the system Q to be classical with respect to some basis $\{|e\rangle\}_{e\in Q}$ then every (k,ϵ)-extractor as in (2) is also a $\big(k+\log(1/\epsilon),2\epsilon\big)$-extractor in the sense of (5). That is, even when the input source is correlated to a classical system Q, every extractor construction still works (nearly) equally well for extracting randomness. However, if Q is quantum no such generic reduction is known and extractor constructions that also work for quantum Q are called quantum-proof.[3] Examples of (approximately) quantum-proof extractors include:

- Spectral (k,ϵ)-extractors are quantum-proof $(k,2\sqrt{\epsilon})$-extractors [8, Theorem 4]. This includes in particular two-universal hashing [26, 32], two-wise independent permutations [30], as well as sample and hash based constructions [18].
- One-bit output (k,ϵ)-extractors are quantum-proof $(k+\log(1/\epsilon),3\sqrt{\epsilon})$-extractors [19, Theorem 1].
- (k,ϵ)-extractors constructed along Trevisan [33] are quantum-proof $\big(k+\log(1/\epsilon),3\sqrt{\epsilon}\big)$-extractors [14, Theorem 4.6] (see also [2]).

We emphasize that all these stability results are specifically tailored proofs that make use of the structure of the particular extractor constructions. In contrast to these findings it was shown by Gavinsky *et al.* [16, Theorem 1] that there exists a valid (though contrived) extractor for which the decrease in the quality of the output randomness has to be at least $\epsilon\mapsto\Omega(m\epsilon)$.[4] As put forward by Ta-Shma [31, Slide 84], this then raises the question if the separation found by Gavinsky *et al.* is maximal, that is:

> *Is every (k,ϵ)-extractor a quantum-proof $\big(O(k+\log(1/\epsilon)),O(m\sqrt{\epsilon})\big)$-extractor or does there exists an extractor that is not quantum-proof with a large separation, say $\epsilon\mapsto(2^m\epsilon)^{\Omega(1)}$?*

We note that such a stability result would make every extractor with reasonable parameters (approximately) quantum-proof. However, for reasons discussed later it is unclear if such a generic quantum-proof reduction is possible and small sets of randomly chosen functions are interesting candidates to study this possibly large classical/quantum separation.

1.3 Our results

- We write the extractor condition (2) as a quadratic optimization program. The optimal value for this program denoted as $\mathrm{C}(\mathrm{Ext},k)$ is the smallest error ϵ such that Ext is a (k,ϵ)-extractor. We then construct a semidefinite programming (SDP) relaxation for this program whose optimal value is denoted $\mathrm{SDP}(\mathrm{Ext},k)$. This program gives an efficiently computable procedure to certify that a family of functions Ext is a (k,ϵ)-extractor for $\epsilon=\mathrm{SDP}(\mathrm{Ext},k)$.

[2] Other notions for weaker quantum adversaries have also been discussed in the literature, e.g., in the bounded storage model (see [14, Section 1] for a detailed overview).

[3] Note that the dimension of Q is unbounded and that it is a priori unclear if there exist any extractor constructions that are quantum-proof (even with arbitrarily worse parameters).

[4] Since the quality of the output randomness of Gavinsky *et al.*'s construction is bad to start with, the decrease $\epsilon\mapsto\Omega(m\epsilon)$ for quantum Q already makes the extractor fail completely in this case.

- We show that this certification procedure gives us much more: it certifies that Ext is a *quantum-proof* $(k, \sqrt{2}\epsilon)$-extractor. Thus, we give a general efficient method for proving that an extractor is quantum-proof. This technique can recover in a unified way many of the currently known methods for constructing quantum-proof extractors. In particular, we can show that constructions based on two-universal hashing [27, 32] are quantum-proof, and that any extractor with entropy deficit $n - k$ or output size m small is quantum-proof [6] (for $m = 1$ this was first shown in [19]). This latter result is a basic building block for showing that Trevisan based extractors are quantum-proof [14].
- We consider the limitations of this SDP relaxation. Even though $\mathrm{SDP}(\mathrm{Ext}, k)$ is a tight bound on $\mathrm{C}(\mathrm{Ext}, k)$ for many extractor constructions, there can be a large gap between these two values. In particular, if $\mathrm{Ext}_{\mathrm{rand}}$ is given by a small number of randomly chosen functions, then $\mathrm{C}(\mathrm{Ext}_{\mathrm{rand}}, k) \ll \mathrm{SDP}(\mathrm{Ext}_{\mathrm{rand}}, k)$. This shows that the method we propose cannot be used to prove that a small set of randomly chosen functions define good extractors. This means that other techniques would be needed to determine whether $\mathrm{Ext}_{\mathrm{rand}}$ is a quantum-proof extractor or not.

2 Preliminaries

2.1 Quantum information

In quantum theory, a system is described by an inner-product space, that we denote here by letters like N, M, Q.[5] Note that we use the same symbol Q to label the system, the corresponding inner-product space and also the dimension of the space. Let $\mathrm{Mat}_Q(S)$ be the vector space of $Q \times Q$ matrices with entries in S. Whenever S is not specified, it is assumed to be the set of complex numbers $\mathbb{C}$, i.e., we write $\mathrm{Mat}_Q(\mathbb{C}) =: \mathrm{Mat}_Q$. The state of a system is defined by a positive semidefinite operator ρ_Q with trace 1 acting on Q. The set of states on system Q is denoted by $\mathcal{S}(Q) \subset \mathrm{Mat}_Q(\mathbb{C})$. The inner-product space of a composite system QN is given by the tensor product of the inner-product spaces $Q \otimes N =: QN$. From a joint state $\rho_{QN} \in \mathcal{S}(QN)$, we can obtain marginals on the system Q by performing a partial trace of the N system $\rho_Q := \mathrm{Tr}_N[\rho_{QN}]$. The state ρ_{QN} of a system QN is called quantum-classical (with respect to some basis) if it can be written as $\rho_{QN} = \sum_x \rho(x) \otimes |x\rangle\langle x|$ for some basis $\{|x\rangle\}$ of N and some positive semidefinite operators $\rho(x)$ acting on Q with $\sum_x \mathrm{Tr}[\rho(x)] = 1$. We denote the maximally mixed state on system N by υ_N.

To measure the distance between two states, we use the trace norm $\|A\|_1 := \mathrm{Tr}[\sqrt{A^*A}]$, where A^* is the conjugate transpose of A. In the special case when A is diagonal, $\|A\|_1$ becomes the familiar ℓ_1 norm of the diagonal entries. Moreover, the Hilbert-Schmidt norm is defined as $\|A\|_2 := \sqrt{\mathrm{Tr}[A^*A]}$, and when A is diagonal this becomes the usual ℓ_2 norm. Another important norm we use is the operator norm, or the largest singular value of A, denoted by $\|A\|_\infty$. When A is diagonal, this corresponds to the familiar ℓ_∞ norm of the diagonal entries. For a probability distribution P_N on the set N, $\|P_N\|_{\ell_\infty}$ corresponds to the optimal probability with which P_N can be guessed successfully. We write

$$H_{\min}(N)_P := -\log \|P_N\|_{\ell_\infty} \ , \tag{6}$$

the min-entropy of P_N. More generally, the conditional min-entropy of N given Q is used to quantify the uncertainty in the system N given the system Q. The conditional min-entropy

[5] In the following all spaces are assumed to be finite-dimensional.

is defined as

$$H_{\min}(N|Q)_\rho := -\log \min_{\sigma_Q \in \mathcal{S}(Q)} \left\| (\mathrm{id}_N \otimes \sigma_Q^{-1/2}) \rho_{NQ} (\mathrm{id}_N \otimes \sigma_Q^{-1/2}) \right\|_\infty , \tag{7}$$

with generalized inverses. Note that in the special case where the system Q is trivial, we have $H_{\min}(N)_\rho = -\log \|\rho_N\|_\infty$.

2.2 Semidefinite programming

Semidefinite programs (SDP) are a large class of optimization problems that can be efficiently solved. Even if one is not explicitly interested in solving it numerically, a semidefinite program often has appealing properties such as strong duality. Semidefinite programming has been extensively used in various contexts in quantum information.

We use a formulation of semidefinite programs sometimes called vector programs. For some fixed values $\alpha_{x,x'}, \beta_{x,x',k}$ and γ_k, the optimization program can be written as follows:

$$\text{maximize} \quad \sum_{x,x'} \alpha_{x,x'} \vec{a}_x \cdot \vec{a}_{x'} \tag{8}$$

$$\text{subject to} \quad \sum_{x,x'} \beta_{x,x',k} \vec{a}_x \cdot \vec{a}_{x'} \leq \gamma_k \quad \text{for all } k \tag{9}$$

Here the optimization is over all vector $\vec{a}_x$ (of arbitrary finite dimension) that satisfy the constraints stated above. Note that we can always assume that the dimension of the vectors $\vec{a}_x$ is bounded by the number of vectors, i.e., the size of the set x runs over.

3 Quadratic programs for randomness extractors

It is useful to see the definition of extractors using the following optimization program:

$$\underline{\text{Error for extractor Ext} = \{f_s\}}$$

$$\mathrm{C}(\mathrm{Ext}, k) := \text{maximize} \quad \frac{1}{D} \sum_{s,y} \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) p(x) \beta_{s,y} \tag{10}$$

$$\text{subject to} \quad 0 \leq p(x) \leq 2^{-k} \tag{11}$$

$$\sum_x p(x) = 1 \tag{12}$$

$$-1 \leq \beta_{s,y} \leq 1 \tag{13}$$

▶ **Definition 1.** Ext is a (k, ϵ)-extractor if and only if $\mathrm{C}(\mathrm{Ext}, k) \leq \epsilon$.

To relate this to the definition given in the introduction, it suffices to observe that the optimal choice for $\beta_{s,y}$ is the sign of $\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) p(x)$ so the objective function becomes $\frac{1}{D} \sum_{s,y} \left| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) p(x) \right|$. The conditions (11) and (12) ensure that the input distribution has min-entropy at least k.

To simplify the program (10) we note that this function is convex in the distribution p and so the maximum is attained in the extreme points of the feasible region. These are simply the distributions that are uniform over a set of size at least 2^k. So we can equivalently write

$$\mathrm{C}(\mathrm{Ext}, k) = \max \left\{ \sum_{s,y} \left| \frac{1}{KD} \sum_{x \in L} \delta_{f_s(x)=y} - \frac{1}{MD} \right| : L \subseteq N, L \geq 2^k \right\} , \tag{14}$$

where again in a slight abuse of notation, we use the letter L for the actual set as well as its size. As the expression being maximized is the ℓ_1 norm between two probability distributions, we can write it as:

$$\mathrm{C}(\mathrm{Ext}, k) = 2 \cdot \max \left\{ \frac{1}{KD} \sum_{x \in L, (y,s) \in R} \delta_{f_s(x)=y} - \frac{R}{MD} : L \subseteq N, L \geq 2^k, R \subseteq M \times D \right\} . \tag{15}$$

This allows us to interpret $\mathrm{C}(\mathrm{Ext}, k)$ in graph-theoretic terms. For that we introduce a bipartite graph with left vertex set N and right vertex set $M \times D$, and there is an edge between vertices x and (y, s) if and only if $f_s(x) = y$. By writing $E(L, R)$ for the set of edges with one endpoint in L and the other endpoint in R, this expression is simply

$$\mathrm{C}(\mathrm{Ext}, k) = 2 \cdot \max \left\{ \frac{E(L,R)}{2^k D} - \frac{R}{MD} : L \subseteq N, L \geq 2^k, R \subseteq M \times D \right\} . \tag{16}$$

Written in this way, we see that the optimization in $\mathrm{C}(\mathrm{Ext}, k)$ is a kind of bipartite densest subgraph problem. Algorithms for a slightly different problem known as the densest K-subgraph problem have been extensively studied, see e.g., [15, 9]. The best known approximation algorithms for this problem achieve a factor of N^α for some constant α, but even ruling out constant factor approximations is only known using quite strong assumptions [1].

We can similarly write a program for the error of Ext against potentially quantum adversaries:

$$\underline{\text{Error for extractor } \mathrm{Ext} = \{f_s\} \text{ against quantum adversaries}}$$

$$\begin{aligned}
\mathrm{Q}(\mathrm{Ext}, k) := \text{maximize} \quad & \frac{1}{D} \sum_{s,y} \sum_{x} \left(\delta_{f_s(x)=y} - \frac{1}{M} \right) \mathrm{Tr}\left[\rho(x) B_{s,y}\right] && (17) \\
\text{subject to} \quad & 0 \leq \rho(x) \leq 2^{-k} \sigma && (18) \\
& \sum_x \mathrm{Tr}[\rho(x)] = 1 && (19) \\
& \mathrm{Tr}[\sigma] = 1 && (20) \\
& \|B_{s,y}\|_\infty \leq 1 && (21)
\end{aligned}$$

Here the maximization is understood over all $\rho(x)$ of arbitrary dimension. Unlike for SDPs for which one can give an upper bound on the dimension of the vector of an optimal solution, no such bound is know in this setting. In fact, we do not even know if the quantity Q is computable.

▶ **Definition 2.** Ext is a quantum-proof (k, ϵ)-extractor if and only if $\mathrm{Q}(\mathrm{Ext}, k) \leq \epsilon$.

To see that this definition coincides with the definition given in the introduction, observe that for fixed $\rho(x)$, the maximum over $B_{s,y}$ of the quantity $\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right) \mathrm{Tr}\left[\rho(x) B_{s,y}\right]$ is $\| \sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right) \rho(x) \|_1$. The constraints on $\rho(x)$ and σ ensure that the state $\sum_x \rho(x) \otimes |x\rangle\langle x|$ has conditional min-entropy at least k.

4 Semidefinite relaxations for randomness extractors

4.1 A relaxation for the extractor condition

Motivated by the fact that the two quantities $\mathrm{C}(\mathrm{Ext}, k)$ and $\mathrm{Q}(\mathrm{Ext}, k)$ are generally difficult to understand, we introduce a SDP that, as we show later, provides a relaxation for both of these quantities. For $\mathrm{Ext} = \{f_s\}_{s \in D}$ and fixed k, we define:

$$\underline{\text{SDP relaxation for error of Ext} = \{f_s\}}$$

$$\text{SDP}(\text{Ext}, k) := \text{maximize} \quad \frac{1}{D}\sum_{s,y,x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x \cdot \vec{b}_{s,y} \tag{22}$$

$$\text{subject to} \quad 0 \leq \vec{a}_x \cdot \vec{a}_{x'} \leq 2^{-k} \cdot q(x) \tag{23}$$

$$q(x) \leq 2^{-k} \tag{24}$$

$$\sum_x q(x) = 1 \tag{25}$$

$$\sum_{x,x'} \vec{a}_x \cdot \vec{a}_{x'} \leq 1 \tag{26}$$

$$\|b_{s,y}\|_2 \leq 1 \tag{27}$$

We maximize over all possible dimensions of the vectors $\vec{a}_x$ and $\vec{b}_x$. Moreover, the Cauchy-Schwarz inequality implies that the optimal choice for $\vec{b}_{s,y}$ is

$$\frac{\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x}{\|\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x\|_2}\,, \tag{28}$$

and thus the objective function of the SDP relaxation becomes

$$\frac{1}{D}\sum_{s,y}\left\|\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x\right\|_2\,, \tag{29}$$

subject to the constraints on the vectors $\vec{a}_x$ stated in (22). By simply plugging $\vec{a}_x = p(x)$, $q(x) = p(x)$ and $\vec{b}_{s,y} = \beta_{s,y}$, we see that this SDP gives an upper bound on the extractor program (10).

▶ **Proposition 3.** *For any* Ext *and* k, $\text{C}(\text{Ext}, k) \leq \text{SDP}(\text{Ext}, k)$. *In other words, if* $\text{SDP}(\text{Ext}, k) \leq \epsilon$, *then* Ext *is a* (k, ϵ)*-extractor.*

This gives a computationally efficient criterion for certifying that an extractor is good. As we show in Section 4.3, this method can certify that many important constructions are good extractors. However, this technique does in general not give a tight characterization of extractors and there can be a large gap between the values $\text{C}(\text{Ext}, k)$ and $\text{SDP}(\text{Ext}, k)$ as we will see in Section 4.4.

4.2 A relaxation for the error against quantum adversaries

A very interesting property about the SDP (22) is that it also gives an upper bound on the error of an extractor against quantum adversaries. This means that if an extractor satisfies the stronger property $\text{SDP}(\text{Ext}, k) \leq \epsilon$ then it is not only a (k, ϵ)-extractor but also a quantum proof $(k, \sqrt{2}\epsilon)$-extractor.

▶ **Theorem 4.** *For any* Ext *and* k, *we have*

$$\text{C}(\text{Ext}, k) \leq \text{Q}(\text{Ext}, k) \leq \sqrt{2} \cdot \text{SDP}(\text{Ext}, k)\,. \tag{30}$$

Proof. Let $\rho = \sum_x \rho(x) \otimes |x\rangle\langle x|$ be a quantum state on QN with $H_{\min}(N|Q)_\rho \geq k$. By the definition of the conditional min-entropy, this implies that there exists $\sigma \in \mathcal{S}(Q)$ such that $\rho(x) \leq 2^{-k}\sigma$ for all $x \in N$. We now define the average state $\bar{\rho} = \sum_x \rho(x)$ and $\omega = \frac{\bar{\rho}+\sigma}{2}$, as well as the vectors $\vec{a}_x$ as the list of entries of the matrix $\frac{1}{\sqrt{2}}\omega^{-1/4}\rho(x)\omega^{-1/4}$. This is so that

we have $\vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{2}\mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')]$. As the trace of the product of two positive semidefinite operators is nonnegative, we have $\vec{a}_x \cdot \vec{a}_{x'} \geq 0$. Moreover, we have

$$\vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{2}\mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')] \leq \frac{1}{2}\mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}2^{-k}\sigma] \tag{31}$$

$$\leq \frac{1}{2}\cdot 2^{-k}\mathrm{Tr}[\omega^{-1/2}\rho(x)\omega^{-1/2}2\omega] \leq 2^{-k}\mathrm{Tr}[\rho(x)] \ . \tag{32}$$

We set $q(x) = \mathrm{Tr}[\rho(x)]$. Note that we have $q(x) = \mathrm{Tr}[\rho(x)] \leq 2^{-k}\mathrm{Tr}[\sigma] = 2^{-k}$ and $\sum_x q(x) \leq 1$. We can also write

$$\sum_{x,x'} \vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{2}\mathrm{Tr}[\omega^{-1/2}\bar{\rho}\omega^{-1/2}\bar{\rho}] \leq \frac{1}{2}\mathrm{Tr}[\omega^{-1/2}\bar{\rho}\omega^{-1/2}2\omega] \leq 1 \ . \tag{33}$$

We now analyze the objective function. We use the following Hölder-type inequality for operators $\|\alpha\beta\gamma\|_1 \leq \||\alpha|^4\|_1^{1/4}\||\beta|^2\|_1^{1/2}\||\gamma|^4\|_1^{1/4}$, see e.g., [10, Corollary IV.2.6]. The error the extractor makes on input ρ is given by

$$\frac{1}{D}\sum_{s,y}\left\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\rho(x)\right\|_1$$

$$\leq \frac{1}{D}\sum_{s,y}\|\omega\|_1^{1/4}\left\|\left(\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\omega^{-1/4}\rho(x)\omega^{-1/4}\right)^2\right\|_1^{1/2}\|\omega\|_1^{1/4} \tag{34}$$

$$= \frac{1}{D}\sum_{s,y}\sqrt{\mathrm{Tr}\left[\sum_{x,x'}\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\left(\delta_{f_s(x')=y}-\frac{1}{M}\right)\omega^{-1/2}\rho(x)\omega^{-1/2}\rho(x')\right]} \tag{35}$$

$$= \frac{1}{D}\sum_{s,y}\sqrt{\sum_{x,x'}\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\left(\delta_{f_s(x')=y}-\frac{1}{M}\right)2\cdot\vec{a}_x\cdot\vec{a}_{x'}} \tag{36}$$

$$= \frac{\sqrt{2}}{D}\sum_{s,y}\left\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\right\|_2 \ . \tag{37}$$

This proves that the error the extractor makes in the presence of quantum adversaries is upper bounded by $\sqrt{2}\cdot\mathrm{SDP}(\mathrm{Ext},k)$. ◀

4.3 Applications

We now give several applications of the SDP relaxation. We show that many results about quantum-proof extractors can be shown with the SDP quantity. First, let us consider general results that do not use the structure of the functions in Ext but simply the extractor's parameters. We know the advantage obtained by a quantum adversary compared to a classical one can by bounded by a function of the number of output bits m or the min-entropy deficit $n-k$ [6] (for $m=1$ this was first shown in [19]). In particular, if m or $n-k$ are small, then the quantum advantage cannot be large. We show that this is actually a property of the SDP.

▶ **Theorem 5.** *For any* Ext *and* k*, we have for any* $\epsilon > 0$*,*

$$\mathrm{SDP}\big(\mathrm{Ext}, k+\log(1/\epsilon)\big) \leq \sqrt{2^m}\sqrt{\mathrm{C}(\mathrm{Ext},k)+\epsilon} \tag{38}$$

$$\mathrm{SDP}(\mathrm{Ext},k) \leq 3K_G 2^{n-k}\mathrm{C}(\mathrm{Ext},k-1) \ , \tag{39}$$

where $K_G \leq 1.8$ *is Grothendieck's constant.*

Proof. As Ext is usually clear from the context, we use $\mathrm{C}(k)$ and $\mathrm{SDP}(k)$ for $\mathrm{C}(\mathrm{Ext}, k)$ and $\mathrm{SDP}(\mathrm{Ext}, k)$. To prove (38), we consider an optimal solution for $\mathrm{SDP}(k + \log(1/\epsilon))$. Define $p(x, x') = \vec{a}_x \cdot \vec{a}_{x'}$, with $\bar{p}(x) = \sum_{x'} p(x, x')$. Now consider the set $S_\epsilon = \{x \in N : \bar{p}(x) \leq \epsilon q(x)\}$. Then $\sum_{x \in S_\epsilon} \bar{p}(x) \leq \epsilon \sum_{x \in S_\epsilon} q(x) \leq \epsilon$. Using the fact that $\vec{a}_x$ define a feasible solution for $\mathrm{SDP}(k + \log(1/\epsilon))$, we have for $x \notin S_\epsilon$, $p(x, x') \leq 2^{-(k+\log(1/\epsilon))} q(x) \leq 2^{-k} \bar{p}(x)$. We can then write using the Cauchy Schwarz inequality,

$$\frac{1}{D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - 2^{-m} \right) \vec{a}_x \right\|_2 \leq \sqrt{\frac{1}{D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - 2^{-m} \right) \vec{a}_x \right\|_2^2} \sqrt{2^m} \, . \tag{40}$$

We now look at the expression $\frac{1}{D} \sum_{s,y} \left\| \sum_x \left(\delta_{f_s(x)=y} - 2^{-m} \right) \vec{a}_x \right\|_2^2$ which equals

$$\frac{1}{D} \sum_{s,y} \sum_{x,x'} \left(\delta_{f_s(x)=y} - 2^{-m} \right) \cdot \left(\delta_{f_s(x')=y} - 2^{-m} \right) p(x, x') \tag{41}$$

$$\leq \frac{1}{D} \sum_{s,y} \sum_x \left| \sum_{x'} \left(\delta_{f_s(x)=y} - 2^{-m} \right) \cdot \left(\delta_{f_s(x')=y} - 2^{-m} \right) p(x, x') \right| \tag{42}$$

$$\leq \frac{1}{D} \sum_{s,y} \sum_x \left| \sum_{x'} \left(\delta_{f_s(x')=y} - 2^{-m} \right) p(x, x') \right| \, . \tag{43}$$

We separate the sum into $x \in S_\epsilon$ and $x \notin S_\epsilon$ and get

$$\frac{1}{D} \sum_{s,y} \sum_x \left| \sum_{x'} \left(\delta_{f_s(x')=y} - 2^{-m} \right) p(x, x') \right| \tag{44}$$

$$= \frac{1}{D} \sum_{s,y} \sum_x \bar{p}(x) \left| \sum_{x'} \left(\delta_{f_s(x')=y} - 2^{-m} \right) \frac{p(x, x')}{\bar{p}(x)} \right| \tag{45}$$

$$= \sum_{x \in S_\epsilon} \bar{p}(x) \frac{1}{D} \sum_{s,y} \left| \sum_{x'} \left(\delta_{f_s(x')=y} - 2^{-m} \right) \frac{p(x, x')}{\bar{p}(x)} \right| \tag{46}$$

$$+ \sum_{x \notin S_\epsilon} \bar{p}(x) \frac{1}{D} \sum_{s,y} \left| \sum_{x'} \left(\delta_{f_s(x')=y} - 2^{-m} \right) \frac{p(x, x')}{\bar{p}(x)} \right| \leq \epsilon + \mathrm{C}(k) \, , \tag{47}$$

which proves (38).

We now prove the inequality (39). For that, we simply upper bound $\mathrm{SDP}(\mathrm{Ext}, k)$ by forgetting several constraints and then apply Grothendieck's inequality (Theorem 9). Observe first that for any feasible vectors $\vec{a}_x$ for the SDP, we have $\|\vec{a}_x\|_2^2 \leq 2^{-k} q(x) \leq 2^{-2k}$.

$$\mathrm{SDP}(\mathrm{Ext}, k) \leq \max \left\{ \frac{1}{D} \sum_{s,y,x} \left(\delta_{f_s(x)=y} - 2^{-m} \right) \vec{a}_x \cdot \vec{b}_{s,y} : \|\vec{a}_x\|_2 \leq 2^{-k}, \|\vec{b}_{s,y}\|_2 \leq 1 \right\} \tag{48}$$

$$\leq K_G \max \left\{ \frac{1}{D} \sum_{s,y,x} \left(\delta_{f_s(x)=y} - 2^{-m} \right) a_x b_{s,y} : |a_x| \leq 2^{-k}, |b_{s,y}| \leq 1 \right\} \tag{49}$$

$$= K_G \max \left\{ \frac{1}{D} \sum_{s,y} \left| \sum_x \left(\delta_{f_s(x)=y} - 2^{-m} \right) a_x \right| : |a_x| \leq 2^{-k} \right\} \, . \tag{50}$$

We partition the set of $x \in N$ into $\{x : a_x \geq 0\}$ and $\{x : a_x < 0\}$ and write

$$\left|\sum_x \left(\delta_{f_s(x)=y} - 2^{-m}\right) a_x\right| \leq \left|\sum_{x:a_x\geq 0} \left(\delta_{f_s(x)=y} - 2^{-m}\right) a_x\right| \tag{51}$$

$$+ \left|\sum_{x:a_x<0} \left(\delta_{f_s(x)=y} - 2^{-m}\right) (-a_x)\right| . \tag{52}$$

Let us write $\alpha_+ := \sum_{x:a_x\geq 0} a_x$. If $\alpha_+ \geq 1$, then we define $p_+(x) = \frac{\max\{a_x,0\}}{\alpha_+}$. Observing that $\alpha_+ \leq 2^{n-k}$, we have

$$\frac{1}{D}\sum_{s,y}\left|\sum_{x:a_x\geq 0} \left(\delta_{f_s(x)=y} - 2^{-m}\right) a_x\right| = \alpha_+ \cdot \frac{1}{D}\sum_{s,y}\left|\sum_x \left(\delta_{f_s(x)=y} - 2^{-m}\right) p_+(x)\right| \tag{53}$$

$$\leq \alpha_+ \mathrm{C}(k + \log(\alpha_+)) \leq 2^{n-k}\mathrm{C}(k) , \tag{54}$$

where we have used the abbreviation $\mathrm{C}(k) = \mathrm{C}(\mathrm{Ext}, k)$. Otherwise (if $\alpha_+ < 1$), we define $p_+(x) = \max\{a_x, 0\} + (1 - \alpha_+)2^{-n}$. We get

$$\frac{1}{D}\sum_{s,y}\left|\sum_{x:a_x\geq 0} \left(\delta_{f_s(x)=y} - 2^{-m}\right) a_x\right| \tag{55}$$

$$= \frac{1}{D}\sum_{s,y}\left|\sum_x \left(\delta_{f_s(x)=y} - 2^{-m}\right) \left(p_+(x) - (1-\alpha_+)2^{-n}\right)\right| \tag{56}$$

$$\leq \frac{1}{D}\sum_{s,y}\left|\sum_x \left(\delta_{f_s(x)=y} - 2^{-m}\right) p_+(x)\right| + (1-\alpha_+)\frac{1}{D}\sum_{s,y}\left|\sum_x \left(\delta_{f_s(x)=y} - \frac{1}{M}\right) 2^{-n}\right| \tag{57}$$

$$\leq \mathrm{C}(k-1) + (1-\alpha_+)\mathrm{C}(n) . \tag{58}$$

With a similar argument for the set $\{x : a_x < 0\}$, we reach the bound

$$\frac{1}{D}\sum_{s,y}\left|\sum_x \left(\delta_{f_s(x)=y} - 2^{-m}\right) a_x\right| \tag{59}$$

$$\leq \max\{2 \cdot 2^{n-k}\mathrm{C}(k), \mathrm{C}(k-1) + \mathrm{C}(n) \tag{60}$$

$$+ 2^{n-k}\mathrm{C}(k), 2\mathrm{C}(k-1) + (1-\alpha_+-\alpha_-)\mathrm{C}(n)\} \leq 3 \cdot 2^{n-k}\mathrm{C}(k-1) . \tag{61}$$

Finally, we get $\mathrm{SDP}(k) \leq 3K_G 2^{n-k}\mathrm{C}(k-1)$. ◀

Some specific constructions are also known to be quantum-proof, in particular constructions based on two-universal hash functions [26, 27, 32]. This type of construction is captured by spectral extractors [8]. For an extractor $\mathrm{Ext} = \{f_s\}_{s\in D}$ we define the linear maps [Ext] and τ that map vectors of dimension N to vectors of dimension DM as follows:

$$[\mathrm{Ext}]\left(\sum_x p(x)|x\rangle\langle x|_N\right) = \frac{1}{D}\cdot\sum_{s,y}\sum_x \delta_{f_s(x)=y} p(x)|y\rangle\langle y|_M \otimes |s\rangle\langle s|_D \tag{62}$$

$$\tau\left(\sum_x p(x)|x\rangle\langle x|_N\right) = \left(\sum_x p(x)\right) v_M \otimes v_D . \tag{63}$$

Note that we used a quantum notation and identified vectors with diagonal matrices. A spectral (k,ϵ)-extractor is then defined via the largest eigenvalue bound

$$\lambda_1\Big([\mathrm{Ext}]^* \cdot [\mathrm{Ext}] - \tau^* \cdot \tau\Big) \leq 2^{k-m-d}\epsilon \ , \tag{64}$$

where $*$ refers to the adjoint of a linear map. We prove next that for spectral extractor, there can be at most a quadratic gap between $\mathrm{C}(\mathrm{Ext},k)$ and $\mathrm{SDP}(\mathrm{Ext},k)$.

▶ **Theorem 6.** *Let* $\mathrm{Ext}_{\mathrm{spec}} = \{f_s\}_{s\in D}$ *be a spectral* (k,ϵ)*-extractor as defined in* (64)*. Then, we have*

$$\mathrm{SDP}(\mathrm{Ext}_{\mathrm{spec}},k) \leq \sqrt{\epsilon} \ . \tag{65}$$

The proof can be found in Appendix B. Another class of extractors that are quantum-proof are Trevisan based constructions [14, 2]. These are particularly important to understand because they are the only known quantum-proof constructions with short seed $d = O(\mathrm{poly}(\log n))$ (cf. the optimal parameters (3)). Trevisan's construction can be thought of as a composition of one-bit output extractors cleverly interleaved by slightly reusing the seed. Specifically, the construction is based on a family of subsets $S_1,\ldots,S_m \subset \{1,\ldots,d\}$ such that for each i we have

$$|S_i| = l \quad \text{and} \quad \sum_{j<i} 2^{|S_i\cap S_j|} \leq r(m-1) \ , \tag{66}$$

for some $r>0$. Such a family $\{S_i\}_{i\in\{1,\ldots,m\}}$ is also called weak (l,r)-design. Now, take a one-bit output extractor $\mathrm{Ext}_{\mathrm{one}} = \{g_t\}_{t\in\{0,1\}^l}$ with $g_t : N \to \{0,1\}$, and a weak (l,r)-design as defined in (66). Trevisan then defines a m-bit output extractor

$$\mathrm{Ext}_{\mathrm{Trev}} = \{f_s\}_{s\in D} \quad \text{with} \quad f_s : N \to M \tag{67}$$

$$f_s(x) := g_{s|S_1}(x) \circ g_{s|S_1}(x) \circ \cdots \circ g_{s|S_m}(x) \ , \tag{68}$$

where $s|S_i$ denotes the l-bits of s that correspond to the position indexed by the set S_i, and $\circ$ means concatenation.[6] The basic idea of the proof is to bound the quality of $\mathrm{Ext}_{\mathrm{Trev}}$ as a function of the quality of $\mathrm{Ext}_{\mathrm{one}}$. Then (using Theorem 5) one can relate the quality of $\mathrm{Ext}_{\mathrm{one}}$ against quantum adversaries to its quality against classical adversaries. We give (in the Appendix) a concise proof of this result using our notation in terms of the quantum program (17).

▶ **Theorem 7.** *Let* $\{S_i\}_{i\in\{1,\ldots,m\}}$ *be a weak* (l,r)*-design as defined in* (66)*, and* $\mathrm{Ext}_{\mathrm{one}} = \{g_t\}_{t\in\{0,1\}^l}$ *be a one-bit output extractor. Then, we have for Trevisan's extractor* $\mathrm{Ext}_{\mathrm{Trev}} = \{f_s\}_{s\in D}$ *as defined in* (67)–(68)*,*

$$\mathrm{Q}(\mathrm{Ext}_{\mathrm{Trev}},k) \leq m \cdot \mathrm{Q}(\mathrm{Ext}_{\mathrm{one}}, k-r(m-1)) \tag{69}$$

$$\leq 2m \cdot \sqrt{\mathrm{C}(\mathrm{Ext}_{\mathrm{one}}, k-r(m-1)-\log(1/\epsilon)) + \epsilon} \ , \tag{70}$$

for any $\epsilon > 0$.

[6] Actual parameters for Trevisan based extractor constructions are, e.g, discussed in detail in [14, Section 5].

4.4 Gap between C and SDP

In this section, we show that there can be a large gap between the value C and SDP. In fact, we show that SDP cannot be used to prove that randomly chosen functions are good randomness extractors. As discussed in (3), random functions are good extractors with essentially optimal parameters. In other words, for a family of functions $\mathrm{Ext}_{\mathrm{rand}} = \{f_s\}_{s\in D}$ chosen at random, we have with very high probability that

$$\mathrm{C}(\mathrm{Ext}_{\mathrm{rand}}, k) \leq \epsilon \quad \text{for} \quad m = k - 2\log(1/\epsilon) - O(1) \tag{71}$$

$$d = \log(n-k) + 2\log(1/\epsilon) + O(1)\ . \tag{72}$$

In contrast to this, we find that the SDP relaxation for random constructions can become very large for sufficiently small min-entropy k.

▶ **Theorem 8.** *Let* $\mathrm{Ext} = \{f_s\}_{s\in D}$ *be a family of functions such that*

$$\gamma_1 \frac{DN^2}{M} \leq \sum_{x,x',s} \delta_{f_s(x)=f_s(x')} \leq \gamma_2 \frac{DN^2}{M}\ , \tag{73}$$

and $k \leq \log\left(\gamma_1 \frac{N}{M}\right)$*. Then, we have*

$$\mathrm{SDP}(\mathrm{Ext}, k) \geq \frac{1}{2}\sqrt{\frac{M}{\gamma_2 D}}\ . \tag{74}$$

When the functions f_s are chosen at random, then the condition (73) is satisfied with very high probability for constant values of γ_1 and γ_2 (see Proposition 11 for a proof). Hence, we find that for instance if $k = n/2$, $m = n/4$ and $d = O(\log n)$, with high probability $\mathrm{SDP}(\mathrm{Ext}_{\mathrm{rand}}, k) \gg 2$, whereas we have with very high probability $\mathrm{C}(\mathrm{Ext}_{\mathrm{rand}}, k) \leq \frac{1}{n}$. As clearly $\mathrm{Q}(\mathrm{Ext}, k) \leq 2$, this also shows that Q can be much smaller than SDP.

Moreover we can show that for Trevisan's extractor, we cannot replace $\mathrm{Q}(\mathrm{Ext}_{\mathrm{Trev}})$ with $\mathrm{SDP}(\mathrm{Ext}_{\mathrm{Trev}}, k)$ in general in Theorem 7. This is because if the one-bit extractors $\{g_t\}$ in Trevisan's construction are chosen at random, then it is possible to show that the condition (73) is satisfied with high probability for constant values of γ_1 and γ_2 (see Proposition 11 for a proof).

Proof of Theorem 8. Use $\vec{a}_x = \alpha^{-1/2} \cdot \sum_{s,y} \delta_{f_s(x)=y} |s\rangle|y\rangle$, $\alpha = \sum_{x,x'} \sum_{s,y} \delta_{f_s(x)=y}\delta_{f_s(x')=y}$. By definition the normalization condition $\sum_{x,x'} \vec{a}_x \cdot \vec{a}_{x'} \leq 1$ is satisfied. Moreover, for any fixed x, x', we have

$$\vec{a}_x \cdot \vec{a}_{x'} = \frac{1}{\alpha} \sum_{s,y} \delta_{f_s(x)=y}\delta_{f_s(x')=y} \leq \frac{D}{\alpha} \leq \frac{1}{\gamma_1}\frac{M}{N^2} \leq \frac{1}{\gamma_1}\frac{M}{N} q(x)\ , \tag{75}$$

where we used the lower bound on γ_1 and we choose $q(x) = 1/N$. Now if $k \leq \log\left(\gamma_1 \frac{N}{M}\right)$, the min-entropy condition for the vectors is satisfied. Now let us analyze the objective function by choosing $\vec{b}_{s,y} = |s\rangle|y\rangle$. We find

$$\frac{1}{D}\sum_{s,y}\sum_{x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\vec{a}_x \cdot \vec{b}_{s,y} = \frac{1}{D}\sum_{s,y}\sum_{x}\left(\delta_{f_s(x)=y} - \frac{1}{M}\right)\alpha^{-1/2}\delta_{f_s(x)=y} \tag{76}$$

$$= \frac{1}{D\alpha^{1/2}}\sum_{s,x}\left(1 - \frac{1}{M}\right) = \frac{N}{\alpha^{1/2}}\left(1 - \frac{1}{M}\right) \geq \frac{1}{2}\sqrt{\frac{M}{\gamma_2 D}}\ , \tag{77}$$

which proves the claim. ◀

5 Discussion

Theorem 8 shows limitations of the SDP relaxation presented here. In fact, even though the error of the extractor $\mathrm{C}(\mathrm{Ext}, k)$ and $\mathrm{Q}(\mathrm{Ext}, k)$ are clearly bounded by 2, the value $\mathrm{SDP}(\mathrm{Ext}, k)$ can be much larger. In [7], we present an improved SDP relaxation that has the property of always being bounded by 2. In addition, we propose a converging hierarchy of SDPs that gives increasingly tight characterizations of quantum-proof extractors.

Acknowledgments. We acknowledge discussions with Matthias Christandl, Fabian Furrer, Patrick Hayden, Christopher Portmann, Renato Renner, Oleg Szehr, Marco Tomamichel, Thomas Vidick, Stephanie Wehner, Reinhard Werner, and Andreas Winter. Part of this work was done while OF and VBS where visiting the Institute for Quantum Information and Matter at Caltech and we would like to thank John Preskill and Thomas Vidick for their hospitality. Most of this work was done while OF was at ETH Zurich supported by the the European Research Council grant No. 258932. Additional funding was provided by the EU under the project "Randomness and Quantum Entanglement" (RAQUEL). VBS was supported by an ETH postdoctoral fellowship.

References

1 Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein. Inapproximability of densest k-subgraph from average case hardness. Manuscript, available at `http://www.csc.kth.se/~rajsekar/papers/dks.pdf`, 2011.

2 Avraham Ben-Aroya and Amnon Ta-Shma. Better short-seed quantum-proof extractors. *Theoretical Computer Science*, 419:17–25, 2012.

3 Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. International Conference on Computers, Systems and Signal Processing*, 1984.

4 Charles H. Bennett, Gilles Brassard, Claude Crépeau, and Ueli M. Maurer. Generalized privacy amplification. *Information Theory, IEEE Transactions on*, 41:1915–1923, 1995.

5 Charles H. Bennett, Gilles Brassard, and Jean-Marc Robert. Privacy amplification by public discussion. *SIAM journal on Computing*, 17:210–229, 1988.

6 Mario Berta, Omar Fawzi, and Volkher B. Scholz. Quantum-proof randomness extractors via operator space theory. *arXiv preprint arXiv:1409.3563*, 2014.

7 Mario Berta, Omar Fawzi, and Volkher B. Scholz. Quantum bilinear programs. *arXiv preprint arXiv:1506.08810*, 2015.

8 Mario Berta, Omar Fawzi, Volkher B. Scholz, and Oleg Szehr. Variations on classical and quantum extractors. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 1474–1478, 2014.

9 Aditya Bhaskara, Venkatesan Guruswami, Moses Charikar, Aravindan Vijayaraghavan, and Yuan Zhou. Polynomial integrality gaps for strong SDP relaxations of Densest k-subgraph. In *Proc. of the 23rd Annual ACM-SIAM Symp. on Discrete Algorithms (SODA'12)*, 2012.

10 R. Bhatia. *Matrix Analysis*. Springer, 1997.

11 Kai-Min Chung, Yaoyun Shi, and Xiaodi Wu. Physical randomness extractors. *arXiv preprint arXiv:1402.4797*, 2014.

12 Roger Colbeck and Adrian Kent. Private randomness expansion with untrusted devices. *Journal of Physics A: Mathematical and Theoretical*, 44:095305, 2011.

13 Matthew Coudron and Henry Yuen. Infinite randomness expansion and amplification with a constant number of devices. In *Proc. of the 46th Annual ACM Symp. on Theory of Computing*, STOC'14, pages 427–436. ACM, 2014.

14 A. De, C. Portmann, T. Vidick, and R. Renner. Trevisan's extractor in the presence of quantum side information. *SIAM Journal on Computing*, 41:915–940, 2012.

15 Uriel Feige and Michael Seltser. On the densest k-subgraph problem. *Algorithmica*, 29:2001, 1997.

16 Dmitry Gavinsky, Julia Kempe, Iordanis Kerenidis, Ran Raz, and Ronald de Wolf. Exponential separations for one-way quantum communication complexity, with applications to cryptography. In *Proc. of the 39th Annual ACM Symp. on Theory of Computing*, STOC'07, pages 516–525. ACM, 2007.

17 Johan Håstad, Russell Impagliazzo, Leonid A Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28:1364–1396, 1999.

18 R. König and R. Renner. Sampling of min-entropy relative to quantum knowledge. *Information Theory, IEEE Transactions on*, 57:4760–4787, 2011.

19 R. König and B. Terhal. The bounded-storage model in the presence of a quantum adversary. *Information Theory, IEEE Transactions on*, 54:749–762, 2008.

20 R. König, S. Wehner, and J. Wullschleger. Unconditional security from noisy quantum storage. *Information Theory, IEEE Transactions on*, 58:1962–1984, 2012.

21 Carl A. Miller and Yaoyun Shi. Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices. *arXiv preprint arXiv:1402.0489*, 2014.

22 Noam Nisan and Amnon Ta-Shma. Extracting randomness: a survey and new constructions. *Journal of Computer and System Sciences*, 58:148–173, 1999.

23 Noam Nisan and David Zuckerman. Randomness is linear in space. *Journal of Computer and System Sciences*, 51:43–52, 1996.

24 Stefano Pironio, Antonio Acín, Serge Massar, A Boyer de La Giroday, Dzimitry N Matsukevich, Peter Maunz, Steven Olmschenk, David Hayes, Le Luo, T Andrew Manning, et al. Random numbers certified by bell's theorem. *Nature*, 464:1021–1024, 2010.

25 J. Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. *SIAM Journal on Discrete Mathematics*, 13:2–24, 2000.

26 Renato Renner. *Security of Quantum Key Distribution.* PhD thesis, ETH Zurich, 2005.

27 Renato Renner and Robert König. Universally composable privacy amplification against quantum adversaries. In Joe Kilian, editor, *Theory of Cryptography*, volume 3378 of *Lecture Notes in Computer Science*, pages 407–425. Springer Berlin Heidelberg, 2005.

28 Ronen Shaltiel. Recent developments in explicit constructions of extractors. *Bulletin of the EATCS*, 77:67–95, 2002.

29 Michael Sipser. Expanders, randomness, or time versus space. *Journal of Computer and System Sciences*, 36:379 – 383, 1988.

30 Oleg Szehr. Decoupling theorems. Master's thesis, ETH Zurich, 2011.

31 Amnon Ta-Shma. Extractors against classical and quantum adversaries. *Tutorial QCrypt*, 2013.

32 M. Tomamichel, C. Schaffner, A. Smith, and R. Renner. Leftover hashing against quantum side information. *Information Theory, IEEE Transactions on*, 57:5524–5535, 2011.

33 Luca Trevisan. Construction of extractors using pseudo-random generators (extended abstract). In *Proc. of the 31st Annual ACM Symposium on Theory of Computing*, STOC'99, pages 141–148. ACM, 1999.

34 Salil P. Vadhan. *Pseudorandomness.* Now Publishers Inc., Hanover, MA, USA, 2012.

35 Umesh Vazirani and Thomas Vidick. Certifiable quantum dice: Or, true random number generation secure against quantum adversaries. In *Proc. of the 44th Annual ACM Symp. on Theory of Computing*, STOC'12, pages 61–76, New York, NY, USA, 2012. ACM.

A Useful Lemmas

▶ **Theorem 9** (Grothendieck's inequality). *For any real matrix* $\{A_{ij}\}$*, we have*

$$\max\left\{\sum_{i,j} A_{ij}\vec{a}_i\cdot\vec{b}_j : \|\vec{a}_i\|_2\leq 1, \|\vec{b}_j\|_2\leq 1\right\} \tag{78}$$

$$\leq K_G\cdot\max\left\{\sum_{i,j} A_{ij}a_ib_j : a_i,b_j\in\mathbb{R}, |a_i|\leq 1, |b_j|\leq 1\right\}\ . \tag{79}$$

▶ **Theorem 10** (Chernoff bound). *Let* $X_i\in\{0,1\}$ *be independent and identically distributed random variables, and* $\mu := \mathbf{E}\{\sum_i X_i\}$*. Then, we have*

$$\mathbf{P}\left\{\sum_i X_i\geq(1+\delta)\mu\right\}\leq\left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu \quad \textit{for any } \delta>0 \tag{80}$$

$$\mathbf{P}\left\{\sum_i X_i\leq(1-\delta)\mu\right\}\leq\left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^\mu \quad \textit{for any } 0<\delta<1\ . \tag{81}$$

B Missing Proofs

Proof of Theorem 6. We start with the expression $\frac{1}{D}\sum_{s,y}\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\|_2$ for the SDP, where the vectors $\vec{a}_x$ fulfill the conditions stated in (22). Using Cauchy-Schwarz, we may bound

$$\frac{1}{D}\sum_{s,y}\left\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\right\|_2\leq\left(\frac{1}{D}\sum_{s,y}\left\|\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\right\|_2^2\right)^{1/2} 2^{m/2}\ . \tag{82}$$

We now take a closer look at the expression in the brackets. Expanding the norm squared gives rise to the expression

$$\frac{1}{D}\sum_{s,y}\left(\sum_x\left(\delta_{f_s(x)=y}-\frac{1}{M}\right)\vec{a}_x\right)\cdot\left(\sum_{x'}\left(\delta_{f_s(x')=y}-\frac{1}{M}\right)\vec{a}_{x'}\right) \tag{83}$$

$$\begin{aligned}
&=\frac{1}{D}\sum_{s,y}\left(\sum_x\delta_{f_s(x)=y}\vec{a}_x\right)\cdot\left(\sum_{x'}\delta_{f_s(x')=y}\vec{a}_{x'}\right)\\
&\quad-\frac{1}{D}\sum_{s,y}\frac{1}{M}\sum_{x,x'}\delta_{f_s(x)=y}\vec{a}_x\cdot\vec{a}_{x'}\\
&\quad-\frac{1}{D}\sum_{s,y}\frac{1}{M}\sum_{x,x'}\delta_{f_s(x')=y}\vec{a}_x\cdot\vec{a}_{x'}\\
&\quad+\frac{1}{D}\frac{1}{M^2}\sum_{s,y}\sum_{x,x'}\vec{a}_x\cdot\vec{a}_{x'}\ .
\end{aligned} \tag{84}$$

Let us examine the cross terms:

$$\frac{1}{D}\sum_{s,y}\frac{1}{M}\sum_{x,x'}\delta_{f_s(x)=y}\vec{a}_x\cdot\vec{a}_{x'}=\frac{1}{D}\sum_s\frac{1}{M}\sum_{x,x'}\vec{a}_x\cdot\vec{a}_{x'}\ , \tag{85}$$

since for each fixed pair $s, x \in D \times N$ there is exactly one $y \in M$ such that $f_s(x) = y$. The second cross term evaluates analogously to the same value, which is also equal to the fourth term in the expansion of the norm, and hence we are left with

$$\frac{1}{D}\sum_{s,y}\left(\sum_x \delta_{f_s(x)=y}\vec{a}_x\right)\cdot\left(\sum_{x'}\delta_{f_s(x')=y}\vec{a}_{x'}\right) - \frac{1}{D}\sum_{s,y}\frac{1}{M}\left(\sum_x \vec{a}_x\right)\cdot\frac{1}{M}\left(\sum_{x'}\vec{a}_{x'}\right) \ . \tag{86}$$

Introducing the maps ψ_s and τ from $\ell_2(N)$ to $\ell_2(M)$,

$$\psi_s : \vec{e}_x \mapsto \sum_y \delta_{f_s(x)=y}\vec{e}_y \quad \text{and} \quad \tau : \vec{e}_x \mapsto \frac{1}{M}\sum_y \vec{e}_y \tag{87}$$

this may be written as

$$\frac{1}{D}\sum_s \psi_s(\vec{a})\cdot\psi_s(\vec{a}) - \tau(\vec{a})\cdot\tau(\vec{a}) \ , \tag{88}$$

where the dot now means taking the scalar product in the Hilbert space $\ell_2(M)\otimes\mathcal{H}$ and we set $\vec{a} = \sum_x \vec{e}_x \otimes \vec{a}_x \in \ell_2(N)\otimes\mathcal{H}$. However, this is up to a factor of $\frac{1}{D}$ exactly the defining expression of a spectral extractor. Hence we may bound

$$\frac{1}{D}\sum_s \psi_s(\vec{a})\cdot\psi_s(\vec{a}) - \tau(\vec{a})\cdot\tau(\vec{a}) \leq 2^k\frac{\epsilon}{M}\|\vec{a}\|^2 \ . \tag{89}$$

The last norm evaluates to

$$\|\vec{a}\|^2 = \sum_x \vec{a}_x\cdot\vec{a}_x \leq 2^{-k}\sum_x q(x) = 2^{-k} \ , \tag{90}$$

and comparison with (82) gives the desired bound. ◀

Proof of Theorem 7. Consider a feasible solution of (17) given by $\rho(x), \sigma, B_{s,y}$ all acting on a Hilbert space Q. The objective function can be written as

$$\begin{aligned}
&\frac{1}{2^d}\sum_{s,y,x}\left(\delta_{f_s(x)=y} - 2^{-m}\right)\mathrm{Tr}[\rho(x)\,B_{s,y}] \\
&= \frac{1}{2^d}\sum_{s,x}\sum_{y\in\{0,1\}}\left(\sum_{t=0}^{m-1}\frac{1}{2^{m-t-1}}\prod_{k=1}^{t+1}\delta_{f_s(x)_k=y_k} - \frac{1}{2^{m-t}}\prod_{k=1}^{t}\delta_{f_s(x)_k=y_k}\right)\mathrm{Tr}[\rho(x)\,B_{s,y}] \qquad (91)\\
&= \sum_{t=0}^{m-1}\frac{1}{2^d}\sum_{s,x}\sum_{y_1,y_2,\ldots y_{t+1}}\prod_{k=1}^{t}\delta_{f_s(x)_k=y_k}\left(\delta_{f_s(x)_{t+1}=y_{t+1}} - \frac{1}{2}\right)\mathrm{Tr}[\rho(x)\,C_{s,y_1,y_2,\ldots,y_{t+1}}] \ , \qquad (92)
\end{aligned}$$

where we defined

$$C_{s,y_1,\ldots,y_t,y_{t+1}} := \frac{1}{2^{m-t-1}}\sum_{y_{t+2},\ldots,y_m\in\{0,1\}} B_{s,y_{t+2},\ldots,y_m} \ . \tag{93}$$

We now start using the particular structure of the extractor in (68). From now, we fix the value of t and the dependence on t of many variables are omitted to lighten the notation. The seed s can be specified by $a = s|S_{t+1} \in \{0,1\}^l$ and $b = s|S^c_{t+1} \in \{0,1\}^{d-l}$ where S^c_{t+1} is

the complement on S_{t+1} in the set $\{1,\dots,d\}$. We will thus interchangeably use s and (a,b). Using this notation with the structure of f_s, we obtain

$$\frac{1}{2^d}\sum_{s,y,x}\left(\delta_{f_s(x)=y}-2^{-m}\right)\mathrm{Tr}[\rho(x)\,B_{s,y}]$$
$$=\sum_{t=0}^{m-1}\frac{1}{2^d}\sum_{\substack{x\\ a\in\{0,1\}^l\\ b\in\{0,1\}^{d-l}}}\sum_{y_1,y_2,\dots y_{t+1}}\delta_{h_{x,b}(a)=y_1\dots y_t}\left(\delta_{g_a(x)=y_{t+1}}-\frac{1}{2}\right)\mathrm{Tr}[\rho(x)\,C_{a,b,y_1,y_2,\dots,y_{t+1}}] \tag{94}$$

$$=\sum_{t=0}^{m-1}\frac{1}{2^l}\sum_{\substack{x\\ a\in\{0,1\}^l}}\sum_{z\in\{0,1\}}\left(\delta_{g_a(x)=z}-\frac{1}{2}\right)\frac{1}{2^{d-l}}\sum_{b\in\{0,1\}^{d-l}}\mathrm{Tr}[\rho(x)\,C_{a,b,h_{x,b}(a),z}] \tag{95}$$

where $h_{x,b}(a)$ represents the first t bits of $f_s(x)$. Note that for a fixed x and b, the outcome of this function only depends on the bits of s that belong to one of the sets $S_1,\dots,S_t$. In particular, the first bit of $h_{x,b}$ only depends on the substring of a corresponding to indices in $S_1\cap S_{t+1}$. Thus, for any x,b, the function $h_{x,b}$ belongs to the family $\mathcal{F}_t$ of functions $h:\{0,1\}^l\to\{0,1\}^t$ for which the j-th bit h^j of h is a function $h^j:\{0,1\}^{S_j\cap S_{t+1}}\to\{0,1\}$. Thus, for any x,b only $\sum_{j=1}^t 2^{|S_j\cap S_{t+1}|}\le r(m-1)$ bits are sufficient to fully describe the function $h_{x,b}$. As a result, $|\mathcal{F}_t|\le 2^{r(m-1)}$.

Let us define new positive operators on a larger $Q\otimes H\otimes G$ system as

$$\hat{\rho}(x):=\frac{1}{2^{d-l}}\sum_{\substack{b\in\{0,1\}^{d-l}\\ h\in\mathcal{F}_t}}\rho(x)\otimes\delta_{h=h_{x,b}}|h\rangle\langle h|_H\otimes|b\rangle\langle b|_G \tag{96}$$

$$\hat{\sigma}:=\frac{1}{|\mathcal{F}_t|2^{d-l}}\sum_{\substack{b\in\{0,1\}^{d-l}\\ h\in\mathcal{F}_t}}\sigma\otimes|h\rangle\langle h|_H\otimes|b\rangle\langle b|_G \tag{97}$$

$$\hat{C}_{a,z}:=\sum_{b,h\in\mathcal{F}_t}C_{a,b,h(a),z}\otimes|h\rangle\langle h|\otimes|b\rangle\langle b|\ . \tag{98}$$

Note that $\hat{\sigma}$ as well as $\sum_x\hat{\rho}(x)$ have unit trace and $\|\hat{C}_{a,z}\|_\infty\le 1$. In addition,

$$\hat{\rho}_x\le\frac{1}{2^{d-l}}\sum_{\substack{b\in\{0,1\}^{d-l}\\ h\in\mathcal{F}_t}}\rho(x)\otimes|h\rangle\langle h|_H\otimes|b\rangle\langle b|_G\le|\mathcal{F}_t|2^{-k}\hat{\sigma}\le 2^{-k+r(m-1)}\hat{\sigma}\ , \tag{99}$$

where we used the fact that $\rho(x)\le 2^{-k}\sigma$. This shows that the newly defined operators $\hat{\rho}(x),\hat{\sigma},\hat{C}_{a,z}$ satisfy the constraints of (17) for the extractor $\mathrm{Ext}_{\mathrm{one}}$ with min-entropy $k-r(m-1)$. Looking at the value of the objective function for this solution, we obtain

$$\frac{1}{2^l}\sum_{a,z,x}\left(\delta_{g_a(x)=z}-\frac{1}{2}\right)\mathrm{Tr}[\hat{\rho}(x)\hat{C}_{a,z}]=\frac{1}{2^l}\sum_{a,z,x}\left(\delta_{g_a(x)=z}-\frac{1}{2}\right)\mathrm{Tr}[\hat{\rho}(x)\hat{C}_{a,z}] \tag{100}$$
$$=\frac{1}{2^l}\sum_{a,z,x}\left(\delta_{g_a(x)=z}-\frac{1}{2}\right)\frac{1}{2^{d-l}}\sum_b\mathrm{Tr}[\rho(x)C_{a,h(a),z}]\ , \tag{101}$$

which is exactly the t-th term in the sum in (95). To relate $\mathrm{Q}(\mathrm{Ext}_{\mathrm{one}},k-r(m-1))$ to $\mathrm{C}(\mathrm{Ext}_{\mathrm{one}},k-r(m-1)-\log(1/\epsilon))+\epsilon$, we use Theorem 4 and Theorem 5. ◀

▶ **Proposition 11.** *Suppose the functions $f_s : N \to M$ from the family $\{f_s\}_{s\in D}$ are chosen at random with $f_s(x)$ and $f_{s'}(x')$ uniformly distributed and independent whenever $x \neq x'$. Then, we have for $N \geq 16$ that*

$$\mathbf{P}\left\{\left|\sum_{x,x',s}\delta_{f_s(x)=f_s(x')} - \left(DN + \frac{DN(N-1)}{M}\right)\right| \geq \frac{1}{2}\frac{DN(N-1)}{M}\right\} \leq \frac{1}{16}\ . \tag{102}$$

This of course includes the case when the functions f_s are chosen uniformly and independently, but also the case of Trevisan's construction where the one-bit extractor is a randomly chosen function.

Proof of Proposition 11. We start by separating the cases $x = x'$ and $x \neq x'$,

$$\sum_{x,x',s}\delta_{f_s(x)=f_s(x')} = DN + \sum_{s,x\neq x'}\delta_{f_s(x)=f_s(x')}\ . \tag{103}$$

We compute the expectation over the choice of f:

$$\mathop{\mathbf{E}}_{f}\left\{\sum_{s,x\neq x'}\delta_{f_s(x)=f_s(x')}\right\} = DN(N-1)\frac{1}{M}\ , \tag{104}$$

simply using the fact then for $x \neq x'$, $f_s(x)$ and $f_s(x')$ are independently chosen. We now would like to show that with high probability this random variable is close to its expectation. For that we compute the second moment

$$\mathop{\mathbf{E}}_{g}\left\{\left(\sum_{s,x\neq x'}\delta_{f_s(x)=f_s(x')}\right)^2\right\} \tag{105}$$

$$= \sum_{s_1,s_2,x_1\neq x_2,x_1'\neq x_2'}\mathbf{P}\left\{f_{s_1}(x_1) = f_{s_1}(x_1'), f_{s_2}(x_2) = f_{s_2}(x_2')\right\} \tag{106}$$

$$= \sum_{s_1,s_2,x_1\neq x_2,x_1'\neq x_2',\{x_1,x_1'\}\neq\{x_2,x_2'\}}\mathbf{P}\left\{f_{s_1}(x_1) = f_{s_1}(x_1'), f_{s_2}(x_2) = f_{s_2}(x_2')\right\} \tag{107}$$

$$+ \sum_{s_1,s_2,x_1\neq x_2,x_1'\neq x_2',\{x_1,x_1'\}=\{x_2,x_2'\}}\mathbf{P}\left\{f_{s_1}(x_1) = f_{s_1}(x_1'), f_{s_1}(x_2) = f_{s_1}(x_2')\right\} \tag{108}$$

$$\leq D^2N(N-1)(N(N-1)-2)\frac{1}{M^2} \tag{109}$$

$$+ 2\sum_{s_1,s_2,x_1\neq x_2}\mathbf{P}\left\{f_{s_1}(x_1) = f_{s_1}(x_1')\right\} \tag{110}$$

$$= D^2N(N-1)(N(N-1)-2)\frac{1}{M^2} + 2D^2N(N-1)\frac{1}{M}\ . \tag{111}$$

As a result the variance is at most

$$\mathbf{Var}\left\{\sum_{s,x\neq x'}\delta_{f_s(x)=f_s(x')}\right\} \tag{112}$$

$$\leq D^2N(N-1)(N(N-1)-2)\frac{1}{M^2} + 2D^2N(N-1)\frac{1}{M} - \left(DN(N-1)\frac{1}{M}\right)^2 \tag{113}$$

$$\leq 2D^2N(N-1)\frac{1}{M}\ . \tag{114}$$

Using Chebyshev's inequality gives with a standard deviation $\sigma \leq \sqrt{2}D\sqrt{N(N-1)/M}$ we have

$$\mathbf{P}\left\{\left|\sum_{s,x\neq x'} \delta_{f_s(x)=f_s(x')} - \frac{DN(N-1)}{M}\right| \geq 4\sigma\right\} \leq \frac{1}{16} \ . \tag{115}$$

But $4\sigma \leq 4\sqrt{2}D\sqrt{N(N-1)/M} \leq \frac{1}{2}\frac{DN(N-1)}{M}$ for $N \geq 16$. ◀